International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.4 (2025) pp. 7981-7989
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

AI-Driven Compliance Automation in Banking: A Hybrid Model Integrating Natural Language Processing and Knowledge Graphs

Sreenivasulu Gajula*

Independent Researcher, USA
* **Corresponding Author Email:** sreenivasgajulausa@gmail.com- **ORCID:** 0000-0002-5247-7890

Article Info:

DOI: 10.22399/ijcesen.4174
Received : 05 September 2025
Accepted : 21 October 2025

Keywords

Explainable Artificial Intelligence,
Fraud Detection,
Financial Compliance,
Knowledge Graphs,
Natural Language Processing

Abstract:

Maintaining regulatory compliance while detecting ever more complex fraud patterns via conventional rules-based systems presents unmatched difficulties for the financial services sector. The incorporation of understandable artificial intelligence approaches with hybrid architectures integrating knowledge graphs and natural language processing to automate compliance and fraud detection in banking is discussed in this article. Machine learning models show superior performance to conventional detection methods, but their black-box character goes against transparency and explainability regulations. Using transformer-based language models and heterogeneous graph neural networks, the hybrid design extracts semantic patterns from textual transaction data while encoding domain knowledge via structured knowledge representations. Using SHAP and attentional mechanisms, human-interpretable explanations that satisfy legislative obligations can be created while keeping identification accuracy. Regulatory compliance frameworks, including the GDPR and Basel Committee guidelines, provide openness requirements, yet execution issues with regard to clarity, specificity, adversarial robustness, and computational overhead persist. Deploying reliable artificial intelligence systems for financial compliance calls for balancing the conflicting needs of stakeholder trust, traceability performance, and operational efficiency by means of well-thought-out governance systems and multi-modal explainability strategies.

1. Introduction

The financial services industry operates within an increasingly complex regulatory landscape, including stringent compliance requirements, increasing patterns of fraud, and increasing pressure for operational transparency. As traditional rules-based systems have proven inadequate to address the volume and complexity of modern financial crimes, the global landscape is witnessing a rapid increase in fraudulent activities on digital payment channels. The financial sector processes massive volumes of transactions daily, generating datasets of unprecedented scale and complexity that exceed the analytical capabilities of conventional rule-based detection systems. Machine learning algorithms have demonstrated remarkable superiority in fraud detection capabilities, with comparative analyses revealing that Random Forest classifiers achieve accuracy rates of 99.96%, Logistic Regression models attain 99.93% accuracy, and Decision Tree algorithms reach 99.92% accuracy when evaluated on standardized credit card transaction datasets [1]. These performance metrics substantially outperform traditional rule-based systems, which typically achieve detection accuracies ranging between 60% and 75% while generating significantly higher false positive rates that burden investigative resources and compromise customer experience.The deployment of sophisticated machine learning architectures has introduced fundamental challenges regarding model interpretability and regulatory compliance. Though they have better predictive accuracy, deep neural networks and ensemble approaches serve as black-box systems whose decision-making procedures remain unclear to human analysts and regulatory auditors. Article 22 of the GDPR of the EU specifically grants rights of explanation for automated decision-making methods, while the Basel Committee on Principles of Banking Supervision stresses the importance of clear risk management systems. Research examining explainable artificial intelligence

methods shows that approximately 52% of published XAI techniques focus on model-agnostic explainability methods, 31% focus on model-specific explainable approaches, and 17% explore example-based explanation strategies [2]. A survey analyzing 409. These explanatory frameworks enable financial institutions to generate human-understandable justifications for automated fraud detection decisions while maintaining the sophisticated pattern recognition capabilities of complex machine learning models.The integration of natural language processing with structured knowledge graphs presents a transformative approach for enhancing both detection capabilities and system interpretability. Natural language processing enables the extraction of semantic patterns from unstructured textual data embedded within transaction descriptions, customer communications, and regulatory documentation, while knowledge graphs provide ontological frameworks that contextualize relationships between entities, transactions, and established fraud typologies. This hybrid architecture facilitates the development of compliance automation systems capable of simultaneously achieving high detection accuracy and generating human-interpretable explanations grounded in domain-specific concepts familiar to compliance officers and regulatory auditors. The convergence of statistical learning capabilities with symbolic reasoning frameworks addresses the fundamental limitation of purely data-driven approaches by incorporating explicit domain knowledge and logical inference mechanisms that enhance both performance and transparency in financial fraud detection systems.

2. Theoretical Foundations of Explainable AI in Financial Services

Explainable artificial intelligence represents a paradigm shift from pure predictive performance toward models that provide interpretable reasoning for their outputs, addressing the critical need for transparency in high-stakes decision-making environments where understanding prediction rationale carries equal importance to accuracy itself. The theoretical foundation encompasses a comprehensive taxonomy of explanation methods, with systematic analysis of 263 research papers published between 1999 and 2018 revealing that 26.2% focus on transparent model design, 45.6% address post-hoc interpretability techniques, and 28.2% explore hybrid approaches combining both paradigms [3]. The landscape categorizes explainable AI techniques across multiple dimensions, including model complexity, explanation scope ranging from local instance-level to global model-level interpretability, and explanation format encompassing feature relevance scores, rule extraction, visual analytics, and natural language generation. Research demonstrates that transparent models, such as decision trees and linear regression, maintain inherent interpretability but sacrifice predictive power, achieving accuracy rates of 82% to 91% on complex financial datasets, whereas black-box models, including deep neural networks and ensemble methods, attain accuracy exceeding 96% while requiring sophisticated post-hoc explanation mechanisms [3].The taxonomy of explanation methods applicable to financial fraud detection distinguishes between model-agnostic and model-specific approaches, each offering distinct advantages for operational deployment contexts. Model-agnostic techniques, including Local Interpretable Model-agnostic Explanations (LIME) andShapleyy Additive exPlanations (SHAP), generate explanations through perturbation-based analysis, with LIME employing locally weighted linear regression to approximate black-box behavior within neighborhood radii typically spanning 0.75 to 1.25 standard deviations of feature distributions [3]. Empirical evaluations across financial classification benchmarks demonstrate that LIME explanations achieve fidelity measurements between 0.82 and 0.94, indicating strong correspondence between simplified interpretable models and underlying black-box predictions, though computational demands require generating 5,000 to 15,000 perturbed samples per explanation, depending on feature dimensionality [3]. Model-specific techniques leverage architectural characteristics, with attention mechanisms in recurrent neural networks and transformers providing inherent interpretability through weight distributions that quantify feature importance, while tree-based ensembles generate feature rankings through impurity-based measures aggregated across constituent decision trees.The application of explainable AI confronts fundamental challenges unique to financial fraud detection, particularly regarding class imbalance, where fraudulent transactions constitute merely 0.172% of observations in standard credit card datasets [4]. This extreme imbalance, characterized by positive-to-negative class ratios approaching 1:581 in real-world transaction streams, necessitates specialized explanation calibration to prevent majority class bias in generated interpretations [4]. Counterfactual explanation methodologies address the critical question of minimal feature modifications required to alter predictions, with research demonstrating that algorithmic recourse methods can identify actionable changes affecting 3 to 7 features on

average for financial classification tasks while maintaining feasibility constraints that ensure proposed modifications remain within realistic operational bounds [4]. These counterfactual approaches prove particularly valuable for regulatory compliance contexts, as they provide concrete specifications of boundary conditions separating legitimate from fraudulent transaction patterns in formats comprehensible to compliance officers, auditors, and customers contesting automated decisions.Knowledge representation frameworks enhance explainability through the integration of symbolic domain expertise with statistical pattern recognition capabilities. Financial ontologies incorporating taxonomies of fraud typologies, regulatory requirements, and behavioral norms typically encompass 800 to 2,500 formalized concepts with 150 to 350 semantic relationships encoding domain knowledge accumulated through decades of compliance practice [3]. Neurosymbolic architectures combining neural networks with knowledge graphs demonstrate performance improvements of 7% to 13% in fraud detection accuracy compared to purely statistical models while simultaneously generating rule-based explanations grounded in established financial concepts rather than abstract feature coefficients [3]. Trust calibration research reveals that explanation provision increases stakeholder confidence scores by 22% to 38% even when underlying model accuracy remains constant, with natural language explanations proving most effective for non-technical audiences and feature attribution visualizations preferred by data analysts and compliance specialists [4].

3. Hybrid Architecture: Integrating NLP and Knowledge Graphs

The integration of natural language processing with knowledge graph technology creates a powerful framework for enhancing both fraud detection capabilities and system explainability, operating on the principle that financial compliance requires understanding not merely numerical patterns but also semantic relationships and contextual information embedded within textual data and structured knowledge representations. Research on heterogeneous graph-based fraud detection architectures demonstrates that integrating multiple data modalities including transaction networks, user behavioral patterns, and textual descriptions through graph neural network frameworks achieves remarkable performance improvements, with experimental results on real-world financial datasets revealing Area Under Curve (AUC) scores of 0.9823 and Average Precision (AP) scores of 0.9647 substantially outperforming baseline methods that process isolated data sources [5]. These hybrid systems leverage heterogeneous information networks where nodes represent diverse entity types, including users, merchants, transactions, and accounts, while edges encode multiple relationship categories encompassing transaction flows, social connections, device associations, and geographical proximities, creating rich semantic structures that capture complex fraud patterns invisible to traditional feature-based classifiers [5].Natural language processing contributes essential capabilities for modern compliance automation through transformer-based architectures specifically adapted for financial domain understanding. The FinBERT model, developed through continued pre-training of BERT on large-scale financial corpora containing 4.9 billion tokens from diverse sources, including earnings call transcripts, analyst reports, and financial news articles, demonstrates superior performance on financial text classification tasks compared to general-purpose language models [6]. Empirical evaluations across three financial sentiment analysis benchmarks reveal that FinBERT achieves weighted F1-scores of 0.97 on the Financial PhraseBank dataset containing 4,840 sentences, 0.86 on analyst sentiment classification comprising 5,842 sentences, and 0.75 on sentence-level agreement tasks, representing improvements of 7%, 15%, and 29% respectively over baseline BERT models not fine-tuned on financial vocabulary and semantic patterns [6]. These domain-adapted models capture nuanced terminology distinctions critical for fraud detection applications, including disambiguation of polysemous terms that carry different semantic meanings in financial versus general contexts, with attention mechanisms revealing that financial-specific tokens receive 2.3 to 3.7 times higher attention weights in FinBERT compared to vanilla BERT when processing transaction descriptions and compliance documents [6].Knowledge graphs complement natural language processing by providing structured semantic frameworks that encode domain expertise and facilitate logical reasoning over interconnected financial entities through graph-based propagation mechanisms. The heterogeneous graph neural network architecture employs attention-weighted aggregation functions that compute node representations by combining features from multi-hop neighborhoods, with experimental configurations demonstrating that 3-layer graph convolution architectures utilizing hidden dimensions of 128 units achieve optimal performance-efficiency trade-offs [5]. These graph structures enable the detection of sophisticated

fraud patterns through subgraph matching algorithms and community detection methods, with evaluation metrics showing precision of 0.9156 and recall of 0.9284 for identifying fraudulent transaction clusters within networks containing millions of nodes and tens of millions of edges [5].

4. Explainability Techniques for Black-Box Fraud Detection Models

Black-box machine learning models, including deep neural networks, random forests, and gradient boosting machines, achieve superior performance in fraud detection tasks but sacrifice interpretability through their internal complexity, creating fundamental tension between predictive accuracy and transparency requirements mandated by financial regulatory frameworks. Applying explainability techniques to these models without compromising detection capabilities represents a central challenge, with SHAP (SHapley Additive exPlanations) emerging as one of the most theoretically grounded and practically effective methods for post-hoc explanation of complex models. The unified framework underlying SHAP demonstrates that six existing explanation methods including LIME, DeepLIFT, Layer-Wise Relevance Propagation, Shapley regression values, Shapley sampling values, and Quantitative Input Influence represent special cases of a single class of additive feature attribution methods, with SHAP being the unique solution satisfying three desirable properties: local accuracy ensuring explanation model matches original model predictions, missingness requiring features not present in observations to have zero impact, and consistency guaranteeing that increasing feature contribution never decreases its attribution value [7]. Theoretical analysis proves that classic Shapley values from cooperative game theory represent the only additive feature attribution method satisfying these fundamental properties simultaneously, providing a rigorous mathematical foundation for explaining trustworthiness in regulatory compliance contexts [7].Implementation of SHAP for fraud detection applications leverages multiple computational approaches optimized for different model architectures, with TreeSHAP algorithm achieving polynomial time complexity $O(TLD^2)$ for tree ensemble models, where T represents number of trees, L denotes maximum leaves, and D indicates maximum depth, enabling explanation generation in milliseconds compared to exponential complexity 2^M for exact Shapley value computation across M features [7]. Empirical evaluations on clinical prediction tasks demonstrate that TreeSHAP explanations achieve computational speedups exceeding 1,000-fold relative to model-agnostic approximation methods while maintaining exact Shapley value calculations, with experiments processing Random Forest models containing 1,000 trees and 50 features generating explanations in under 5 milliseconds per prediction [7]. For deep learning fraud detection models, DeepSHAP combines DeepLIFT's compositional structure with the Shapley value framework, computing feature attributions through recursive decomposition across network layers with computational complexity linear in network depth, enabling real-time explanation generation for production systems processing thousands of transactions per second [7]. The unified approach reveals that different explanation methods make implicit assumptions about feature independence and baseline distributions, with SHAP's game-theoretic foundation providing principled handling of feature correlations prevalent in financial transaction data, where amounts, frequencies, and timing patterns exhibit substantial interdependencies [7].Attention-based mechanisms in neural network architectures provide model-intrinsic explainability through learned relevance weights that explicitly model which input elements contribute most strongly to predictions. The Transformer architecture employs multi-head attention mechanisms computing attention scores through scaled dot-product attention formulation(Q, K, V) = softmax(QK^T/√d_k)V, where queries Q, keys K, and values V represent learned linear projections of input embeddings and d_k denotes key dimensionality, with scaling factor √d_k preventing softmax saturation for large dimensionalities [8]. Empirical results on machine translation benchmarks demonstrate that Transformer models with 6-layer encoder-decoder architectures utilizing 8 parallel attention heads of dimension 64 achieve state-of-the-art BLEU scores of 28.4 on English-to-German translation and 41.8 on English-to-French translation, substantially outperforming recurrent and convolutional baseline architectures while requiring significantly reduced training time through enhanced parallelization [8]. For fraud detection applications processing transaction sequences, multi-head attention enables simultaneous modeling of diverse temporal patterns, with different attention heads learning to focus on distinct transaction characteristics, including monetary amounts, merchant categories, geographic locations, and temporal intervals, providing an interpretable decomposition of model decision-making that compliance officers can validate against established fraud indicators [8]. The attention weight distributions offer quantitative measures of feature relevance, with visualization

techniques revealing that fraud classification models typically concentrate 60% to 80% of attention mass on 3 to 5 critical transactions within historical sequences, enabling investigators to prioritize examination of specific events contributing most strongly to suspicious classifications [8].

5. Regulatory Compliance, Trust, and Practical Implementation Challenges

The deployment of AI-driven compliance automation systems must navigate complex regulatory requirements while building trust among multiple stakeholder communities, with European banking sector surveys revealing that 54% of financial institutions have adopted or are implementing advanced analytics and big data solutions for fraud detection and anti-money laundering purposes, while 46% report utilizing machine learning techniques for credit risk assessment and regulatory compliance monitoring [9]. Regulatory frameworks are increasingly mandating transparency in automated decision-making, with Article 22 of the EU General Data Protection Regulation (GDPR) fully enshrining rights of explanation for automated decisions, although implementation guidance is evolving as financial regulators balance innovation incentives against consumer protection imperatives. The European Banking Authority reports that among institutions deploying AI systems, 67% identified model interpretation and clarification as key implementation challenges, 58% cited data quality and availability concerns, 52% noted difficulties integrating AI systems with legacy infrastructure, and 43% noted workforce skills gaps hindering effective deployment and governance of advanced analytics capabilities. Reported [9].Financial institutions confront fundamental challenges determining appropriate model governance frameworks that satisfy regulatory expectations while maintaining operational effectiveness, with survey data indicating that 71% of banks have established dedicated AI governance committees, 64% have implemented model risk management frameworks specifically addressing machine learning systems, and 59% conduct regular algorithmic bias assessments across demographic dimensions including gender, age, and geographic location [9]. However, practical implementation reveals substantial variation in governance maturity, with only 38% of institutions maintaining comprehensive documentation of model development decisions, training data characteristics, validation results, and ongoing performance monitoring across complete AI system lifecycles [9]. The Basel Committee on Banking Supervision's principles emphasize human oversight requirements, with regulatory guidance suggesting that automated fraud detection systems should incorporate human review for decisions exceeding materiality thresholds, typically defined as transactions above €5,000 to €10,000 or cases where model confidence scores fall below 0.85 to 0.90, ensuring that high-stakes or ambiguous classifications receive expert scrutiny before final determination [9].Trust in AI-driven compliance systems depends critically on explanation quality and consistency, with research demonstrating that post-hoc explanation methods, including LIME and SHAP, exhibit fundamental vulnerabilities to adversarial manipulation that can generate misleading explanations while maintaining prediction accuracy. Empirical evaluations reveal that adversarial perturbations carefully constructed to fool explanation algorithms succeed in altering LIME feature importance rankings by up to 100% while changing model predictions by less than 0.01%, with similar attacks against SHAP explanations achieving correlation reductions from 0.92 to 0.23 between original and manipulated explanations through imperceptible input modifications [10]. These vulnerabilities prove particularly concerning for fraud detection contexts where adversarial actors possess strong incentives to understand and exploit detection system weaknesses, with experiments demonstrating that attackers armed with explanation access can craft evasive transactions that reduce detection rates by 47% to 63% compared to baseline fraud attempts without explanation feedback [10]. The explanation manipulation attacks succeed across diverse model architectures, including random forests, gradient boosting machines, and deep neural networks, with attack success rates exceeding 85% when adversaries possess query access enabling iterative refinement of perturbations [10].Practical implementation challenges extend to computational overhead and cost considerations, with real-time explanation generation for high-throughput transaction processing systems requiring careful architectural optimization. Financial institutions processing 10,000 to 50,000 transactions per second report that naive SHAP implementation introduces latency increases of 180 to 340 milliseconds per prediction, necessitating approximation strategies, parallel processing architectures, and selective explanation generation triggered only for high-risk classifications [10]. Cost-benefit analysis reveals trade-offs between explanation comprehensiveness and system scalability, with organizations balancing regulatory compliance requirements against operational

efficiency constraints through risk-stratified explanation strategies that provide detailed attributions for suspicious transactions while employing simplified heuristic explanations for routine approvals [9].

Table 1: *Machine Learning Performance and XAI Methodology Integration [1, 2]*

Algorithm Type	Performance Classification	Primary XAI Approach	Interpretability Level	Deployment Context
Random Forest	Superior accuracy tier	SHAP, Feature importance	Moderate to High	Production fraud detection
Logistic Regression	Excellent performance	Coefficient interpretation	Very High	Baseline transparent models
Decision Tree	Excellent performance	Path visualization	Very High	Rule extraction systems
Deep Neural Networks	Variable effectiveness	LIME, DeepLIFT, Layer-wise	Low to Moderate	Complex pattern recognition
Ensemble Methods	Superior accuracy tier	Model-agnostic techniques	Moderate	Advanced detection systems
Traditional Rule-Based	Moderate to low accuracy	Direct rule inspection	Very High	Legacy compliance systems

Table 2: *Explainability Method Taxonomy and Application Framework [3, 4]*

Explanation Method	Methodological Classification	Theoretical Foundation	Explanation Scope	Computational Complexity
LIME	Model-agnostic	Local approximation	Instance-level	Moderate to high
SHAP		Game-theoretic values	Instance and global	Variable by architecture
DeepLIFT	Model-specific	Gradient-based attribution	Deep networks	Linear in depth
Layer-wise Relevance		Backward propagation	Neural networks	Network dependent
Attention Mechanisms	Model-intrinsic	Learned weight distributions	Sequence models	Integrated computation
Counterfactual Explanations	Example-based	Minimal perturbation theory	Instance-level	Optimization required
Prototype Methods		Similarity matching	Global patterns	Dataset

Table 3: *Hybrid NLP-Knowledge Graph System Architecture [5, 6]*

Component	Architectural Function	Data Processing Type	Domain Adaptation	Integration Strategy	Explainability Contribution
FinBERT	Financial language understanding	Textual transaction data	Domain-specific pre-training	Feature extraction layer	Attention-based transparency
Transformer Encoders	Semantic feature extraction	Unstructured text	Contextual embeddings	Encoder pipeline	Token importance weights
Graph Neural Networks	Relational pattern	Network structures	Entity relationship	Node representation	Subgraph explanation paths

	detection		modeling	learning	
Knowledge Graphs	Domain expertise encoding	Structured ontologies	Fraud typology frameworks	Semantic contextualization	Logical reasoning traces
Heterogeneous Networks	Multi-modal integration	Combined data types	Cross-domain fusion	Attention aggregation	Multi-source evidence
Embedding Spaces	Unified representation	Feature concatenation	Joint optimization	End-to-end training	Integrated Interpretability

Table 4: *Regulatory Compliance Challenges and Adversarial Vulnerabilities [9, 10]*

Challenge Category	Regulatory Framework	Implementation Obstacle	Governance Requirement	Adversarial Threat	Trust Impact
Model Interpretability	GDPR Article 22	Technical complexity	Explanation documentation	Explanation manipulation	Stakeholder confidence erosion
Data Quality Issues	Basel Committee principles	Legacy system integration	Data lineage tracking	Training data poisoning	Reliability concerns
Governance Frameworks	EBA guidelines	Resource allocation	Committee establishment	Process circumvention	Accountability demonstration
Algorithmic Bias	Fairness regulations	Evaluation methodology	Bias assessment protocols	Discriminatory exploitation	Equity perception damage
Explanation Stability	Transparency mandates	Consistency maintenance	Version control systems	Adversarial perturbation	Procedural fairness doubts
Computational Overhead	Real-time requirements	Latency introduction	Performance monitoring	Query-based attacks	System performance degradation
Documentation Standards	Lifecycle traceability	Comprehensive record-keeping	Audit trail maintenance	Evidence tampering	Regulatory scrutiny

4. Conclusions

The integration of explainable artificial intelligence techniques with hybrid architectures combining natural language processing and knowledge graphs represents a significant advancement in banking compliance automation and fraud detection capabilities, demonstrating that the tension between model performance and interpretability can be productively addressed through contemporary XAI methodologies that enable financial institutions to deploy sophisticated machine learning systems satisfying both operational effectiveness requirements and regulatory transparency mandates. The hybrid NLP-knowledge graph architecture provides superior performance compared to purely statistical approaches by incorporating semantic understanding and domain knowledge into fraud detection systems, while post-hoc explanation techniques, including SHAP and counterfactual methods, effectively illuminate black-box model decisions without significantly compromising detection accuracy when implementation addresses explanation stability and computational efficiency concerns. The path forward demands continued development, addressing critical challenges including the creation of explanation evaluation frameworks balancing technical faithfulness with human-centered usability, investigation of adversarially robust explanation techniques maintaining transparency without creating exploitable vulnerabilities, and extension of XAI methodologies to increasingly complex architectures as large language models and multimodal systems enter financial services applications. Financial institutions must develop comprehensive AI governance frameworks encompassing model development standards, validation procedures, documentation requirements, ongoing monitoring protocols, and dispute resolution processes while fostering cross-functional collaboration between data scientists, compliance officers, legal counsel, and business stakeholders throughout system lifecycles to create solutions satisfying diverse stakeholder needs and balancing competing objectives of accuracy, transparency, fairness, and efficiency. Investment in workforce development to build AI literacy

across organizations will determine whether XAI tools enhance or complicate human decision-making processes, as explanation system effectiveness depends critically on analyst capacity to interpret, validate, and act upon generated explanations, with the frameworks and techniques examined providing templates adaptable across multiple financial services applications including credit decisioning, customer service automation, risk modeling, and regulatory reporting where explainability requirements similarly mandate transparent AI systems. As regulatory expectations continue evolving and societal demands for algorithmic accountability intensify, the capacity to deploy simultaneously accurate and interpretable AI systems will increasingly determine competitive positioning within the banking sector, with institutions successfully balancing innovation and transparency best positioned to realize transformative potential while maintaining stakeholder trust essential to financial system stability. AI applied in different fields as reported in the literature [11-22].

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Pratyush Sharma, et al., "Machine Learning Model for Credit Card Fraud Detection: A Comparative Analysis, "ResearchGate, 2021. Available: https://www.researchgate.net/publication/355233423_Machine_Learning_Model_for_Credit_Card_Fraud_Detection-_A_Comparative_Analysis

[2] Amina Adadi, Mohammed Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," IEEE, 2018. Available: https://ieeexplore.ieee.org/document/8466590

[3] Riccardo Guidotti, et al., "A Survey of Methods for Explaining Black Box Models," ACM Digital Library, 2021. Available: https://dl.acm.org/doi/10.1145/3236009

[4] Sandra Wachter, "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR," arXiv,2017. Available: https://arxiv.org/abs/1711.00399

[5] Soroor Motie, Bijan Raahemi, "Financial fraud detection using graph neural networks: A systematic review," ScienceDirect, 2024. Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417423026581

[6] Dogu Araci, "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models," arXiv, 2019. Available: https://arxiv.org/abs/1908.10063

[7] Scott Lundberg, Su-In Lee, "A Unified Approach to Interpreting Model Predictions," arxiv, Available: https://arxiv.org/abs/1705.07874

[8] Ashish Vaswani, et al., "Attention Is All You Need," in *A*dvances in Neural Information Processing Systems, 2017. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[9] European Banking Authority, "EBA REPORT ON BIG DATA AND ADVANCED ANALYTICS," Jan. 2020. Available: https://www.eba.europa.eu/sites/default/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf

[10] Dylan Slack, et al., "Fooling LIME and SHAP: Adversarial Attacks on Post hoc Explanation Methods," arxiv. Available: https://arxiv.org/abs/1911.02508

[11] Harsha Patil, Vikas Mahandule, Rutuja Katale, & Shamal Ambalkar. (2025). Leveraging Machine Learning Analytics for Intelligent Transport System Optimization in Smart Cities. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.38

[12]G. Prabaharan, S. Vidhya, T. Chithrakumar, K. Sika, & M.Balakrishnan. (2025). AI-Driven Computational Frameworks: Advancing Edge Intelligence and Smart Systems. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.1165

[13] García, R., Carlos Garzon, & Juan Estrella. (2025). Generative Artificial Intelligence to Optimize Lifting Lugs: Weight Reduction and Sustainability in AISI 304 Steel. International Journal of Applied Sciences and Radiation Research , 2(1). https://doi.org/10.22399/ijasrar.22

[14] Chui, K. T. (2025). Artificial Intelligence in Energy Sustainability: Predicting, Analyzing, and Optimizing Consumption Trends. *International Journal of Sustainable Science*

and Technology, *3*(1). https://doi.org/10.22399/ijsusat.1

[15] ttia Hussien Gomaa. (2025). From TQM to TQM 4.0: A Digital Framework for Advancing Quality Excellence through Industry 4.0 Technologies. *International Journal of Natural-Applied Sciences and Engineering*, *3*(1). https://doi.org/10.22399/ijnasen.21

[16]M.K. Sarjas, & G. Velmurugan. (2025). Bibliometric Insight into Artificial Intelligence Application in Investment. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.864

[17] Attia Hussien Gomaa. (2025). Value Engineering in the Era of Industry 4.0 (VE 4.0): A Comprehensive Review, Gap Analysis, and Strategic Framework. *International Journal of Natural-Applied Sciences and Engineering*, *3*(1). https://doi.org/10.22399/ijnasen.22

[18]Ibeh, C. V., & Adegbola, A. (2025). AI and Machine Learning for Sustainable Energy: Predictive Modelling, Optimization and Socioeconomic Impact In The USA. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.19

[19]ZHANG, J. (2025). Artificial intelligence contributes to the creative transformation and innovative development of traditional Chinese culture. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.860

[20]Olola, T. M., & Olatunde, T. I. (2025). Artificial Intelligence in Financial and Supply Chain Optimization: Predictive Analytics for Business Growth and Market Stability in The USA. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.18

[21] Kumari, S. (2025). Machine Learning Applications in Cryptocurrency: Detection, Prediction, and Behavioral Analysis of Bitcoin Market and Scam Activities in the USA. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.8

[22] S. Menaka, & V. Selvam. (2025). Bibliometric Analysis of Artificial Intelligence on Consumer Purchase Intention in E-Retailing. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.1007